Generalisation of 1:50k roads centrelines from 1:10k road polygons in an automated workflow

Vincent van Altena[1], Marc Post[1], Ron Nijhuis[1], Jantien Stoter[1,2]

[1]Kadaster, Apeldoorn, The Netherlands, firstname.secondname@kadaster.nl
[2]Delft University of Technology, The Netherlands, j.e.stoter@tudelft.nl

1. Introduction

This paper reports about on going developments in the map production process of the Dutch Kadaster (who also hold the mapping agency).
In March 2013, the Dutch Kadaster will produce the 1:50k map series fully automatically from 1:10k data for the first time. From that moment on, this new product will replace the existing 1:50k map. It is the first time that a complete topographic map is generalised with no human interaction and that the resulting map is good enough to replace the existing map.
The implementation is a 100% automated generalisation workflow that produces a countrywide map series at scale 1:50k from 1:10k data within 50 hours. Therefore a 1:50k map delivery is foreseen with every 1:10k data update, i.e. five times a year.
The research that achieved the fully automated generalisation workflow was explained in several publications (Stoter et al, 2009; 2011; 2012). The automated workflow for the generalisation of 1:50k maps from TOP10NL data is currently being extended to produce the 1:100k map series.
This workshop paper focuses on the generalisation problem for which one third of the sub models were developed: generalisation of the road network. The Dutch data models at 1:10k and 1:50k have two properties that characterise this process. At first 1:10k data contains road polygons for every road wider than 2 meters that represent the real widths of the roads, while 1:50k maps contain road centerlines that cover complete roads and that have an attribute indicating the width of the road. The second property that characterises the generalisation of roads is that data at all scale form a planar partition, i.e. no gaps or overlap are allowed. Therefore every polygon that is removed should be replaced by something else.

2. The generalisation problem of roads

For the generalisation of the 1:50k roads from 1:10k data (TOP10NL), four major (generalisation) operations need to be applied. These steps seem straightforward. However, as will be explained in Section 3, considerable effort was paid (with self-developed models) to cover the exceptional cases. The four main generalisation steps for roads are:

1. Replace the 1:10k road polygons with centrelines. This is done by using the TOP10NL centrelines as source data. In TOP10NL this road network is stored additionally to the road polygons that are used for the map. In TOP10NL every road wider than 2 meter is represented by two geometries: a polygon (for the map) and a centre line per lane for the road network. In contrast, road centrelines at

scale 1:50k cover complete roads (see Figure 1). Therefore the road TOP10NL centrelines are collapsed with the “merge divided roads” tool of ArcGIS. In this operation, roads with the same code are collapsed.

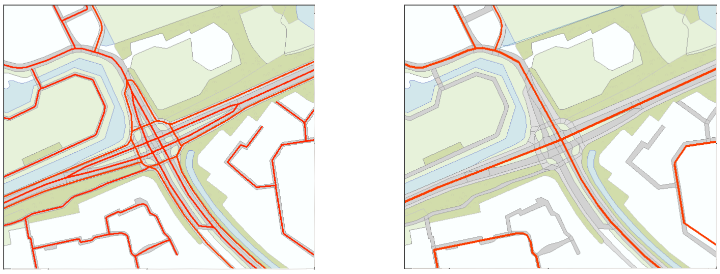

Figure 1 Centrelines in TOP10NL (left) and TOP50 (right), both projected on TOP10NL polygons

2. Extend land use parcels to road centrelines by a self-developed algorithm. Due to replacing of road polygons with road centrelines, gaps exist in a small buffer around the road centrelines. The original road areas need to be assigned to the neighbouring area objects to assure a complete coverage of topographic classes. This is done by extending land use parcels to the new centrelines. To accomplish this, the land use parcel polygons are converted to boundaries, where all boundaries that share a segment with the boundary of a road polygon are firstly removed. Then the remaining land use boundaries are extended. The extended land use boundaries together with the road centrelines (result of first step) represent the new 1:50k land use polygons.
3. Prune the road network, see also Chaudhry and Mackaness (2006), Thom (2007) and Thompson and Richardson (1999). This is not straightforward, specifically because TOP10NL data does not contain many attributes for pruning. The pruning of the road network is done with the ‘thin road network algorithm’ available in ArcGIS. This algorithm retains connectivity and the general character while using a hierarchy for the relative importance and the minimum length of a road. For the time being, we did not further investigate a self-developed algorithm based on road continuity. Figure 2 shows the road network before and after pruning.

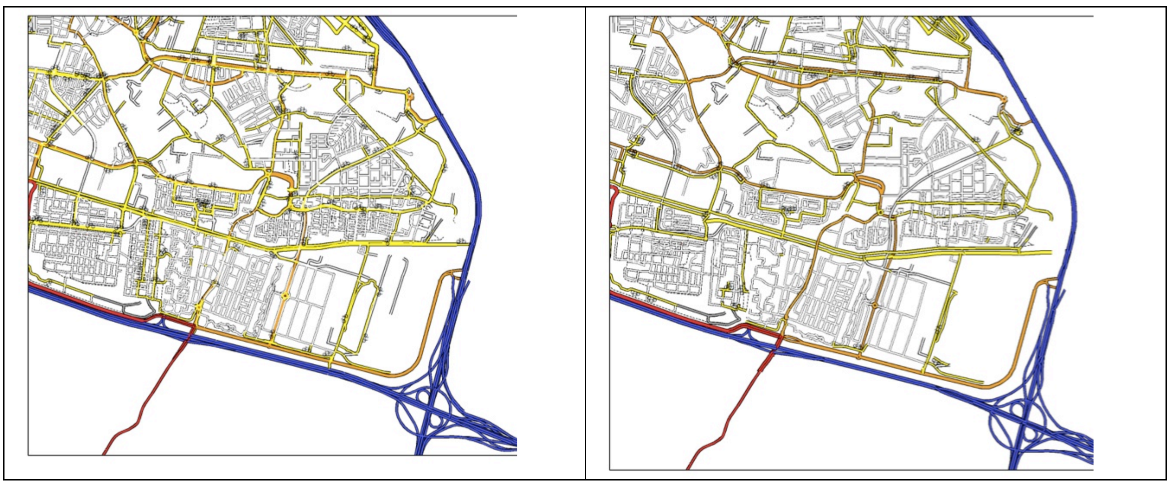

Figure 2 Thinning of the road network before (left) and after (right) generalisation (displayed with same symbol size and at same scale)

In the displacement process at the end of the generalisation workflow, symbolised features are displaced. In this displacement process, roads get highest priority, i.e. they are never displaced in favor of keeping the location of other object types. This is because the (main) road network is used to partition the country to be able to perform countrywide generalisation. The (main) roads are often physical boundaries in the real world and therefore hardly any object is divided by these linear boundaries. Using these real-world boundaries as partition-boundaries makes it less complicated to keep the topology at partition-boundaries. However, if road density is still too high, the roads may get displaced because of each other. This should be avoided as much as possible.

3. The specific cases

The generic approach, as presented in the previous section, generalises input roads (1:10k) containing both road polygons (inclusive adjacent road polygons) and road centrelines to non adjacent road lines (this can still be parallel roads) and prunes the network in a next step, see Figure 3.

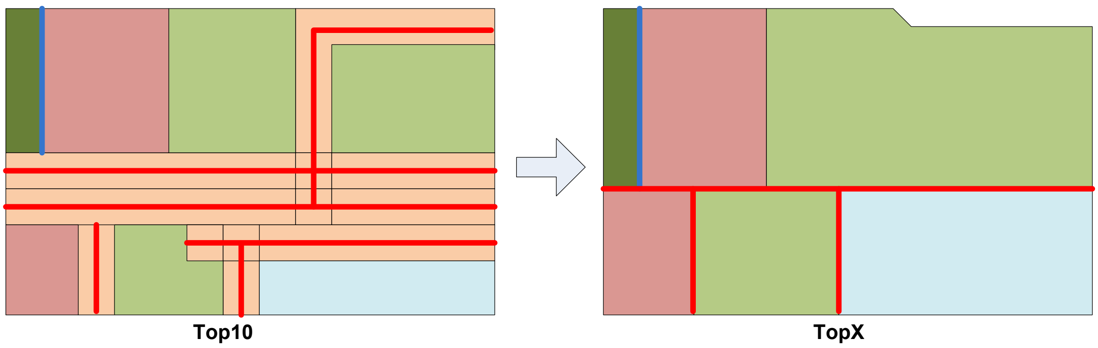

Figure 3 Generalisation of roads: source data (left) en target data (right)

This section presents (a selection of) special cases that could not be generalised with the generic approach (only). Most of these specific cases became clear after having evaluated intermediate results.
The solution of each case consists of two steps. At first the case (i.e. specific context) needs to be identified in an automated manner. Secondly the appropriate generalisation operation needs to be performed. For each of the selected exceptional cases, the specific problem, automated detection and solution is presented below.

__Case 1:__
Problem: Cycle paths that are (partly) adjacent to roads should become part of the main road in 1:50k. Otherwise the displacement pushes the two roads away from each other in the displacement process (Figure 4).
Approach:
1. Detection: Select polygons representing cycle paths that touch road-polygons
2. Solution:
a. Cycle paths adjacent to roads are deleted; the others (i.e. free cycle paths) are maintained.
b. The centreline representing the 1:50k cycle path is extended to the adjacent road to retain the connectivity of the roads. In addition, the extension assures the planar partition.

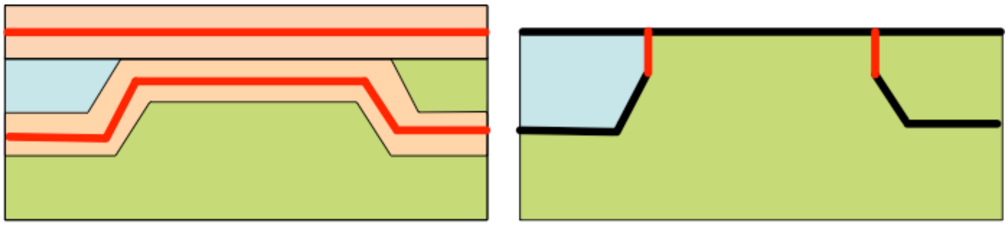

Figure 4 Generalisation of cycle tracks that are adjacent to roads

Case 2:
Problem: Two land use polygons enclosed by roads with different encodings cannot be identified as such by the standard procedure (see section 2), because boundaries of land uses are removed when they intersecting with boundaries of road polygons (Figure 5).
Approach:
1. Detection: Select polygons of roads that touch each other and check whether land use is enclosed.
2. Solution: Select the touching boundary of the road polygons. Create a line perpendicular to the centre point of the touching boundary. This line is connected on either side to the both adjacent road lines. In this way there is a new boundary created between the two land use polygons, see figure below.

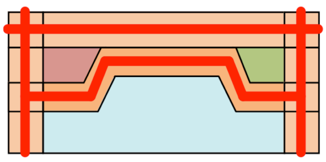	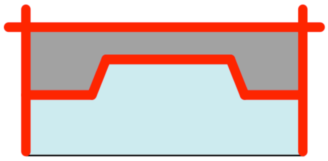	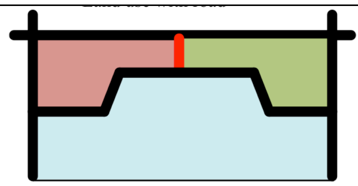
Land use with road polygons	*Unwanted land use with road centre lines*	*Wanted generalisation of land use enclosed by roads*

Figure 5 Generalisation of land use enclosed by roads

Case 3:
Problem: Adjacent roads with different code should not be merged but the less important one should be deleted (Figure 6).
Approach:
1. Detection: Select polygons representing roads that share a boundary with other road-polygons, with different codes.
2. Solution: select the boundary between the differently encoded road and the road polygon. Remove the centre line belonging to the adjacent differently encoded road polygon.
After removing the adjacent roads, the land use areas are extended to the remaining road lines.

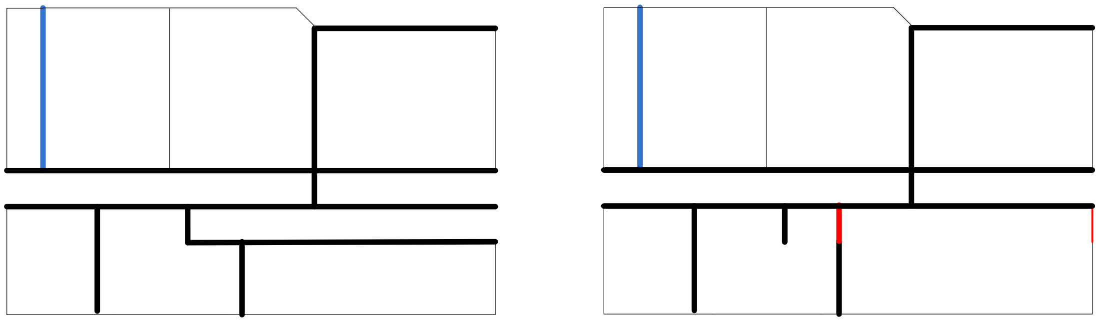

Figure 6 Merging of roads should not process roads with different codes

Case 4:

Problem: In TOP10NL, every road consists of multiple geometries: polygon and centre line covering exactly the complete polygon (not more, i.e. they do not intersect the outer boundary). In situations of road polygons that touch in the corners this leads to unwanted cul-de-sacs, as can be seen in the figure below. In addition, if the specific road is removed in the pruning process, these centerlines are to be used as new boundaries of the adjacent land use polygons and should therefore be connected as a boundary line (Figure 7).

Approach:

1. Detection: Find all dead ends of the road lines that are touching the boundary of the road polygons
2. Solution: Extend the road lines to each other. If the roads are pruned later on in the process, the extended line segment is appended to the feature class of (soft) land use boundaries.

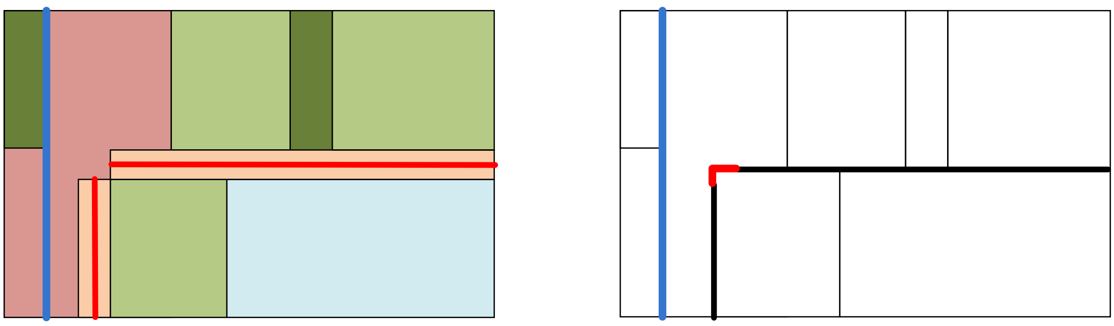

Figure 7 Generalisation of roads with incomplete network geometries

Case 5:

Problem:

Small water areas (or other polygonal object) enclosed by roads disappear because only polygon boundaries are considered in the displacement process (step 4 in section 2), not the polygons themselves. The enclosed area will become invisible, because its boundaries are road lines (with a wide representation), i.e. these areas are covered by the road symbols. The generic displacement approach assures sufficient space *between* objects; but does not provide sufficient space *within* the enclosed objects themselves. See Figure 8.

Approach:

1. Detection: select water (or other polygonal objects) totally enclosed by roads.
2. Solution (see steps in Figure 9): Save information of "disappearing" polygon in an attribute point and reassign this information to the rebuilt polygons at the end of the generalization workflow. To create space in between the two roads, the displacement

takes into account the symbolisation of roads at 1:50k. This assures that the roads are not covering each other. To prevent that they are covering the area in between, the representation of roads is exaggerated in the process, while the common symbolisation is used. The result is a wider displacement.

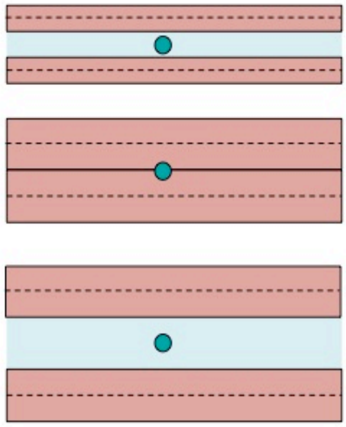

Figure 8 Generalisation of water enclosed by parallel roads

Case 6:

Problem:

The visualisation of parallel roads, which are the results of merging individual lanes (i.e. highways) are still representing a too wide area in the 1:50k dataset (Figure 9).

Approach:

1. Detection: Select road polygons that are adjacent to each other and have the same road classification.

2. Solution: Repeat the merge-road centreline operation. During the merging of the divided roads, the connectivity is retained and attributes are copied. After collapsing the roads, the land use boundaries are extended to the new road line (again). So part of the process as presented in section 2 is repeated.

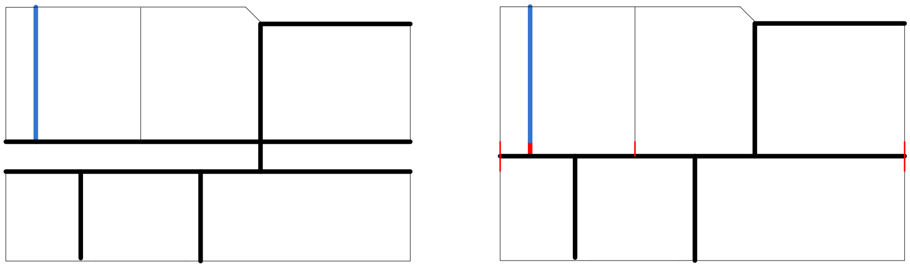

Figure 9 Symbolised paralell roads (left after merging of roads) cover a too wide area in the map and should therefore be merged

Case 7:

Problem:

Highway exits are not encoded in TOP10NL. Because these geometries are encoded as highways, they are included in the merge dual road operations. This gives unwanted results at exits.

Approach: The TOP10NL data model is adjusted and the attribute of the existence of an exit is added. This attribute is filled for the complete data set in one effort, which makes it possible to exclude these road segments from the merging process in the future.

4. Conclusions and further research

This paper reports on a automated map production process currently practised by the Dutch Kadaster. In this process a countrywide 1:50k map series is generalised fully automatically from 1:10k data.
Although a generic approach solves most of the cases, the challenge of the implementation of a full production line is solving exceptional cases. This paper presents a selection of exceptional cases that have to be addressed in the generalisation of road networks. Indeed, our experiments showed that the difficulty of automated generalisation lies not in the automation of the large part (say 80%) but in solving the exceptions.
On-going work is addressing a few other cases to improve next versions of the 1:50k map. For roads this is for example the generalisation of multi-level infrastructural crossing. TOP10NL data contains a road network in 2D and therefore the generalisation of junctions at unequal heights does not always correctly reflect that one road is crossing the other, specifically at complicated junctions. Generalisation of these fly-over crossings is being improved by taking the (relative) heights into account. This information is available in TOP10NL data.
Other on-going work is extending the generalisation workflow to produce the 1:100k map series fully automatically from TOP10NL data. Most submodels developed in the 1:10k to 1:50k generalisation can be reused here (with different parameter values). However, a few issues are more challenging. For example the used road-thinning algorithm is sufficient to generalise most (artificially constructed) waterways for 1:50k representations. However sometimes narrow waterways (under 40m) need to be merged into lines (i.e. collapsed) which brings extra challenges, apart from the lack of a thorough centreline tool. What to do for example if only a small part of a watercourse is (a little) too narrow: ignore the small section, collapse only the narrow section or widen this section to the minimum size?

Acknowledgements
This research is supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs (project code: 11300).

References

Chaudhry, O. Z. and Mackaness, W.A. (2006) Rural and Urban Road Network Generalization Deriving 1:250,000 From 1:1250. In International Cartographic Conference pp. 10p. 10-16 July Spain

Stoter, JE, Smaalen, J van, Nijhuis, R, Dortland, A, Bulder, J & Bruns, B (2012). Fully automated generalisation of topographic data in current geo-information environments. In S Zlatanova, H Ledoux, EM Fendel & M Rumor (Eds.), Urban and Regional Data Management - UDMS Annual 2012 (pp. 111-121). Leiden: CRC Press - Taylor & Francis Group.

Stoter, J.E. and Ron Nijhuis Marc Post Vincent Van Altena Jan Bulder Ben Bruns John Van Smaalen (2011), Feasibility study on an automated generalisation production line for multiscale topographic products, 14th Workshop of the ICA commission on Generalisation and Multiple Representation, June, 2011, Paris

Stoter, J.E., J. van Smaalen, N. Bakker, P. Hardy (2009). Specifying map requirements for automated generalisation of topographic data, The Cartographic Journal Vol. 46 No. 3 pp. 214–227 August 2009

Thom, S. (2007). Automatic resolution of road network conflicts using displacement algorithms orchestrated by software agents, 10th ICA Workshop on Generalization and Multiple Representation, Moscow.

Thompson, R. C., and Richardson, D. E. (1999). The Good Continuation Principle of Perceptual Organization applied to the Generalization of Road Networks, Proceedings of the ICA 19th International Cartographic Conference (Ottawa, 14-21 August), 1215-1223